Comments on:

Lower Willamette Group
Portland Harbor Numerical Modeling

Comments by:

David A. Jay
Professor
Department of
Civil and Environmental Engineering
Portland State University

December 2014

Introduction

The objectives of this review are to: a) provide context for evaluation of numerical modeling of Portland Harbor physical processes and contaminant transport, and b) critique the numerical modeling carried out so far. It is important to provide contextual information and discuss issues that to date have not considered. The numerical modeling critique will focus on modules and outputs relating to circulation, sediment transport, bed layering, and contaminant fate and transport. An analysis of system hydrology is included in Appendix I, while Appendix II defines sediment loading of Portland Harbor from the Willamette River.

I. Context

Conceptual model of physical and sedimentary processes: Determining what processes may disturb the bed and/or drive import or export of sediments is vital to choosing hydrodynamic and sediment transport modeling approaches and to defining boundary conditions for model runs. Thus, definition of a conceptual site model or CSM it vital. Physical aspects of the CSM are covered in the "*Revised Phase 2 Recalibration Results: Hydrodynamic Sedimentation Modeling for Lower Willamette River*," Section 1.5 [West Consultants and Tetra Tech, Inc., 2009].

1. Important conceptual issues discussed in the CSM: the above physical CSM raises several important points:
   a) *Complex circulation patterns associated with Multnomah Slough*: The physical CSM correctly emphasizes the complexity of circulation processes occurring between the Columbia River (Rm-0) and the Willamette River end of Multnomah Channel at Rm-3[1], and that deposition occurs in this part of the Willamette River channel. Given a potential for accumulation of contaminants related to complex circulation in this area and low computational cost, it is inexplicable that the decision was made not to extend the Columbia River grid domain at least to a point downstream of the intersection of the Columbia River with Multnomah Channel near St Helens, so that these circulation processes could be better modeled.[2]
   b) *Importance of bedload*: The physical CSM also emphasizes the importance of bedload transport, mentioning that about half the sediment transport from upstream into the Study Area (Rm-1.0 to 11.8) occurs via bedload, and arguing that a downstream decrease in bedload is important to deposition in the Study Area. It is then extremely puzzling then, that the decision was made in the RI/FS (Appendix La) not to model bedload. This appears to be a major omission that calls into question all results based on the sediment transport modeling.

[1] River-mile (Rm) designations apply to the Willamette River, unless otherwise noted.
[2] Correct representation of flood events actually requires a yet larger grid, as discussed below.

c) *Mix of deposition and erosion*: The physical CSM indicates that patterns of erosion and deposition in Portland Harbor are complex in time and space. Considerable anthropogenic disturbance (e.g., due to propwash) occurs, particularly in shallow areas. It is stated that a major flood can cause up to 1 m of erosion in channel areas, with smaller effects in shallow areas. However, the model results on which this statement is based are poorly verified, as discussed below, and major uncertainties remain. Because both the annual cycle and flood events typically include erosion and deposition, the details of sediment supply and transport matter. Not enough thought has been given to the importance of and the variable nature of extreme events. As discussed below, the 100-year and 500 year flood events have also been greatly underestimated, and the modeling described in the RI/FS Appendix La is not consistent with the physical CSM, in that bedload is not modeled, despite its importance.

2. <u>Conceptual issues not discussed in the CSM</u>: There are several important conceptual issues not discussed, or inadequately discussed, in the physical CSM. These include:

   a) *Maximum bedstresses*: The highest bedstresses with potential for erosion and export are likely associated with the rising phase of Willamette River floods that occur during periods of relatively low Columbia River flow. If Columbia River flows are rising or high, then bedstresses will be reduced due to the deeper depths in Portland Harbor and/or inflows from the Columbia.

   b) *Rain-on-snow winter flooding*: These floods are particularly important, because flows rise rapidly and the supply of fine sediment from upriver is large, leading to the potential for erosion (and export) followed by deposition. The Willamette usually rises faster than the Columbia, but the erosion potential of some winter floods is probably reduced by Columbia River flow management that causes artificially high water levels. Moreover, the fine sediment supply associated with rain-on-snow floods may differ from that which occurs under other conditions.

   c) *The importance of tides*: The Willamette is an estuary (tidal to Oregon City), though an usual one because of the absence of salinity intrusion. Tides are somewhat larger than in the adjacent Columbia River. They are largest during periods of low river flow, but are damped greatly by high flows in either of both rivers. Tidal currents are likely fairly weak, but that does not mean that they have no role in sediment transport. There is also a flood-dominant asymmetry in the tidal currents (at least most of the year) which may facilitate landward transport. It should be demonstrated that the numerical circulation model EFDC is correctly reproducing tides and tidal currents.

d) *Baroclinic processes*: Although there is no salinity intrusion into Portland Harbor, this does not necessarily mean that there are no stratified flow processes. The Willamette may be either warmer or cooler than the Columbia, even on a monthly averaged basis [Rodriguez et al., 2001]. When the Columbia is cooler, this density difference, though small, will favor landward near-bed transport, especially in combination with the asymmetric tides. There may be other periods, however, when the Columbia River is warmer, because of the storage of heat in Columbia River reservoirs. This could lead to a tendency for seaward near-bed currents. Further, vertical profile data in Rodriguez et al. [2001] indicate that vertical temperature differences in summer may reach 3 °C in profiles. These differences may arise from either surface heating or temperature gradients between the Willamette and Columbia Rivers. Also, water temperatures in Portland Harbor may rise by as much as 5-7 °C at the beginning of a winter flood. This inevitably leads to horizontal density gradients that may drive baroclinic circulation processes. These stratified flow processes should be assessed through observations, to see if they need to be modeled, but there appear to be insufficient data available to accomplish this.

e) *Importance of extreme events*: Model results suggest that Portland Harbor undergoes annual cycles of deposition and erosion. Deposition and erosion also occur during storms, and correctly estimating the depth of storm erosion is clearly critical. Unfortunately, the flood with the best available data [December 1964; Waananen et al., 1970, 1971] has not been modeled. Instead, the 1996 storm, for which almost no data are available, was modeled instead. Also, the rather complex repertoire of combined Willamette and Columbia River floods has not been considered. For example during the 1964 flood, the Columbia (at Vancouver) and Willamette Rivers flows peaked almost simultaneously on the morning or 25 December which likely increased the water levels and reduced the currents in Portland Habor on that day. During the 1996 flood, in contrast, flows peaked at Bonneville Dam the day before and the day after the peak flow at Portland, likely due to flood control efforts. However, significant rises do not occur on the Columbia during all Willamette River flood events, which will alter the mean slope, water levels and currents in Portland Harbor. These issues require much more consideration than they have been given, if deposition and erosion are to be understood and correctly modeled.

f) *Flood styles*: Willamette River floods occur in winter. Jay & Naik [2011] argued that there are three styles of winter floods in the combined Columbia and Willamette River system: i) Western sub-basin floods with extensive snowmelt, like February 1881, 1890, 1894 and 1955; ii) combined Interior and Western sub-basin floods like January 1881, 1964 and 1996; and iii) Western sub-basin floods floods without extensive snowmelt like 1909, and 1923.[3] These events likely have different signatures in terms of their influence on Portland Harbor sediment supply, erosion and deposition. Given the importance of floods in Portland Harbor sediment dynamics, this issue should be considered in the conceptual model and explored via numerical modeling.

g) *Extreme bedstresses*: Erosion is caused by bedstress, not high water levels, though the two may be associated. The maximum bedstress in Portland Harbor probably occurs, not at the peak of a flood (when Columbia River flows are also likely to be high), but when water levels in Portland Harbor are rising rapidly from a low level at the onset of a flood. Thus, there is more to designing a 100 year flood scenario than just defining the maximum discharge. The entire flow history and interactions with the Columbia River both matter, and several types of events need to be considered.

h) *System sedimentology*: It is unclear whether the gravel and coarse sand found in Portland Harbor is contemporary or relict. It source is also unclear – the Clackamas River, Johnson Creek and bank erosion are the most likely sources. It is important to understand the sedimentology to: a) understand likely extreme events, and b) determine whether the numerical modeling is transporting sediment correctly.

i) *Climate change impacts*: Modeling is carried out for a 45 year period from 1979. However, no consideration has been given to how climate change has altered and is altering the Columbia and Willamette Rivers hydrological cycles, and sediment supply. Future floods may (or may not) be different from those over the last century, and the impacts of such changes on Portland Harbor contaminant stability need to be considered.

3. <u>Recommendations</u>: The Lower Willamette River is an estuary, with typically complex physical processes. A physical CSM has to explain the overall physical circulation and transport pathways of the Study Area. It should also encompass the range of physical processes, extreme events and climate change impacts that are likely to affect the Study Area. What are the predominant circulation patterns? How important are tides and baroclinic circulation processes, quantitatively? To say deposition occurs during low flow periods and erosion during high flow events is not enough. How, exactly, do these processes occur? How far upstream does tidal current reversal occur under different combinations of Columbia and Willamette River flows? How important are tidal currents to

[3] The Interior sub-basin is the Columbia River watershed east of the Cascade Mountains. The Western sub-basin includes all tributaries west of the Cascades, including the Willamette and other near-coastal rivers.

deposition of sediment during low-flow periods. Are there locations in Portland Harbor where suspended concentrations are typically high and material is trapped on the bed, or is there a general downstream increase in concentrations? These descriptions should, moreover, be based on a mix of data analyses and modeling results. Unfortunately basic observational data have not been collected in the system, making it difficult to understand the processes.

In terms of models, it is particularly vital to determine whether baroclinic circulation processes significantly affect Portland Harbor. If they do, the present generation of 2D (horizontal) models are likely inappropriate. Even if a 2D hydrodynamic model is appropriate, model verification has also not been adequate, and this problem should be corrected. In summary, it is necessary to actually understand the circulation and sediment/contaminant transport dynamics of Portland Harbor, if remedial alternatives are to be realistically evaluated.

II. Adequacy and use of oceanographic data

1. <u>Data available and missing</u>: The data available for numerical circulation model calibration are:
   a. *Water level time series*: Tidal data from three long-term US Geological Survey (USGS) gauges and several shorter National Oceanographic and Atmospheric Administration (NOAA) gauges. These USGS gauges are located at the mouth of Columbia Slough, at Oregon City, and at the Morrison Street Bridge. The Morrison Street water-level record dates back to 1876 (under other auspices), and could be used to assess historical changes in the system. Given the general paucity of data, it is odd that Columbia Slough and Oregon City data sets have not been used at all. In addition, the model grid fails to encompass nearby tide gauges in Vancouver and Longview. These are needed for validation of the EFDC circulation modeling.
   b. *Current time series*: USGS side-looking acoustic Doppler current profiler (ADCP) data from the Morrison Street Bridge gauging station since 2003. This is the only time series of velocity data known to me; it should be used for model calibration and validation. This ADCP also has acoustic backscatter data that could be used (with calibration) to assess water column suspended sediment concentrations. Additional, multi-year ADCP time-series with acoustic backscatter are also needed.
   c. *Lateral and alongchannel velocity transect data*: Miscellaneous ADCP transect data have been collected for short current periods, but these are not adequate for model

   calibration. In particular no data are available for the extreme high flow conditions that are likely to move sediment.

   d. *Profile data*: There are no vertical density profile data in the system from which to judge the importance of baroclinic processes, beyond those presented in Rodriquez et al. [2001]. Simultaneous time-series measurements from multiple depths at convenient locations (e.g., the Morrison Street and St Johns Bridges) should be made to resolve this issue. The time series should cover at least one annual cycle and be continued until a significant flood event has been captured. If vertical and horizontal temperature variations are significant enough to merit inclusion in modeling, then time series at varies locations along the lower Willamette River from Kelley Point to Oregon City will be needed.

   e. *Time-series turbidity data*: The only time series data available, from the Morrison Street Bridge (since 2009), have not been used for model calibration and validation; they should be. It would be sensible to carry out multiple level turbidity sampling along with the temperature sampling described in the previous paragraph. Turbidity time series should be calibrated with water level sampling.

2. <u>Recommendation</u>s: The available data set is not sufficient to allow a thorough calibration and validation of hydrodynamic and sediment transport models of Portland Harbor. Without proper calibration and validation of these models, it is not possible to accurately model contaminant transport, or to defend an analysis of remediation alternatives based thereon. Time series ADCP data should be collected at key locations in the system, time series of temperature and turbidity data should be collected (and for the turbidity data, calibrated with water samples), to determine whether baroclinic processes are important and to better understand suspended sediment distribution and transport. Other measurements, particularly time series measurements covering floods, should be made. Long-term moorings are needed for this purpose.

III. Understanding historical trends and climate change

1. <u>Historical trends and climate change</u>: It is important to understand the historic trajectory of a system for several reasons. First, it is useful to understand the transport conditions under which pollutants accumulated. Do those conditions still pertain (so that contaminants trapped on the bed are likely to remain), or has the system changed in a fundamental way, perhaps making contaminants more mobile? Also, what are the ongoing trends that will influence future transport conditions? Further, how will ongoing climate change alter the system? There has been a long-term decrease in water levels in the Columbia River at Vancouver due to decreased bed friction (better channel alignment), a

deeper channel, and an excess of sand removal over supply [Jay et al., 2011; Templeton and Jay, 2013]. Similar conclusions apply also to Portland Harbor, though we have not published these results as yet. Further, there have been climate-related decreases in in average Willamette and Columbia River flow that are not explained by irrigation withdrawal [Naik and Jay, 2011]. While the total flow of both systems has decreased due to climate change and irrigation withdrawal, this does not imply that the potential for large floods has decreased, because a warmer climate allows for more rapid snow melt. For example, warmer conditions bring the possibility of a late winter Columbia River flood that combines the characteristics of winter and spring floods; i.e., a rapid rise due to rain-on-snow leading to melting of most of the snow pack. Such a flood could be larger than any historical event since ca. 1800. Finally, the Port of Portland has indicated that dredging of Portland Harbor to 43 ft is needed for the commercial viability of the harbor. How will this affect water levels, currents and sediment dynamics in the system?

2. <u>Recommendations</u>: As noted above, climate change impacts and possible human alterations need to be considered in analyses of the system and design of remedial alternatives. The historical trajectory of the system needs to be understood in order to realistically predict future trends that may contaminant dynamics.

IV. Analyses of Model Bias and Uncertainty

1. <u>Climate modeling and superfund modeling, similarities</u>: Regional climate/ hydrologic modeling is similar to the modeling done for a superfund site in several respects:
   a) *Models are chained*: In regional climate/hydrologic modeling, a global climate model drives a regional climate mode, which then drives a hydrologic model. For the Portland Harbor Superfund site, a hydrodynamic model drives a sediment transport model, which drives fate and transport modeling. Human risk models are then developed based on the fate and transport modeling.
   b) *Prediction*: There is a need to project decades into the future, requiring long model runs, so that there is a necessary trade-off between model resolution and computational effort.
   c) *Complexity*: Climate models, like the sediment/fate and transport models, have many sub-modules that parameterize complex, poorly understood processes, amplifying the possibilities for random errors and bias.
   d) *Error propagation*: There is a need to assess the propagation of errors and uncertainty from one model to the next. This is vital in the superfund case, because the remedial alternatives cost very different amounts. Accordingly, there is a

need to determine whether the results of these alternatives can be meaningfully distinguished. This can only be done if propagation of errors from one model to the next is assessed.

2. <u>An inadequate framework for analysis or uncertainty and bias has been used</u>: The RI/FS framework for bias and uncertainty analysis is inadequate. Several aspects of uncertainty analysis have been neglected:

   a) *Errors in individual models*: Errors have not been correctly assessed within individual models. For example, the most basic aspects of performance of the EFDC circulation model have not been assessed – statistics on reproduction of tides, water levels, tidal currents and mean currents. Further, no assessment of errors associated with the neglect of baroclinic processes in the EFDC circulation module has been made, even with respect to the circulation processes. The effects of limited grid resolution and extent have not been evaluated, even in the circulation modeling.

   b) *Accumulation of errors in chained models*: Models are inevitably uncertain and incomplete, and errors propagate from one model to the next. When models are chained so that a circulation model drives a sediment transport module, which then forces a fate and transport model, errors accumulate, as suggested schematically in Figure 1. That is, the environmental data input at each step are uncertain or in error – the boundary sediment input approach, as discussed in Appendix B, is a good example of this. The limited grid resolution and extent used in the circulation modeling affect further calculation in the sediment transport and fate and transport modeling. Further uncertainty occurs in each model due to algorithmic approximations and errors, and uncertain parameters. The use of a vertically integrated model, for example, eliminates baroclinic processes that may be important and requires use of uncertain approximations in formulating sediment deposition from the water column in the sediment transport model. The complex formulation of fine sediment behavior in the water column and bed, which involves numerous poorly known parameters, is a good example of parametric uncertainty.

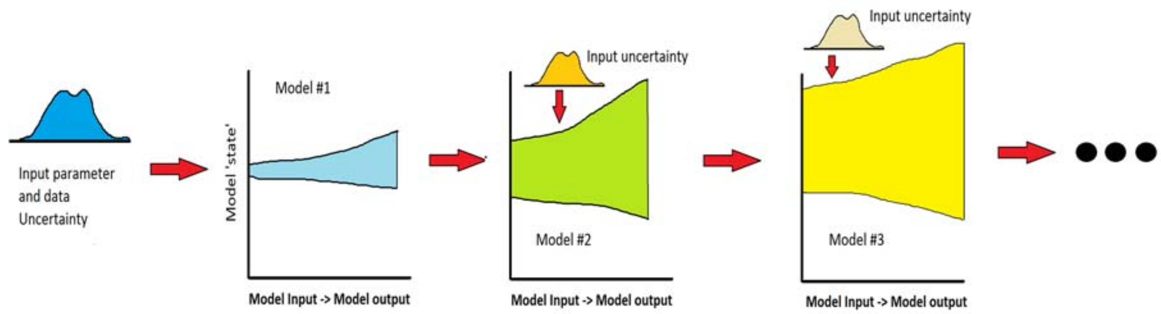

Figure 1: Model uncertainty increases as models are chained. Environmental input data to the first model and to all subsequent models are uncertain and possibly biased. Also, models are put together in sequence, the uncertainty in the prior model(s) must be considered in estimating the uncertainty at any level in the modeling hierarchy. Results from a nonlinear model also become more uncertain over time as errors accumulate (which is why weather can only be predicted a short time in advance), but that is not illustrated here. (Image courtesy of Dr. Stefan Talke, PSU.)

c) *Accumulation of errors over time*: Model errors accumulate over time, because the processes modeled are chaotic; i.e., small errors associated with boundary conditions, paramterizations, algorithmic representations, and limited grid resolution and extent accumulate over time, within a single model and between models. Also, unknown future environmental conditions and climate change affect predictions. This is why, for example, hurricane modeling in done using an ensemble approach, as suggested by Figure 2. Thus, an ensemble approach to long term prediction is needed to reduce uncertainty from all causes..

d) *Evaluation of model bias*: The statistics of model uncertainties and errors can be analyzed by an ensemble approach. However, evaluation of systematic model bias generally requires a multi-modeling approach. The basic idea is expressed conceptually in Figure 3. Models are both uncertain and biased. Ensemble modeling reduces uncertainty due to the effects of random errors. Multi-modeling is needed to reduce errors related to systematic biases. In the present case, 2D and 3D implementations of EFDC could be used for this purpose.

Figure 2: Ensemble forecasts are needed to deal with temporal accumulation of errors in models and also with the effects of uncertain future environmental conditions, as suggested by the diversity of forecasts for the path of Hurricane Sandy a few days before landfall. Reliance on a single forecast would most likely have resulted in drastically incorrect predictions. (Image http://images.huffingtonpost.com/2012-10-25-image001.jpg.)

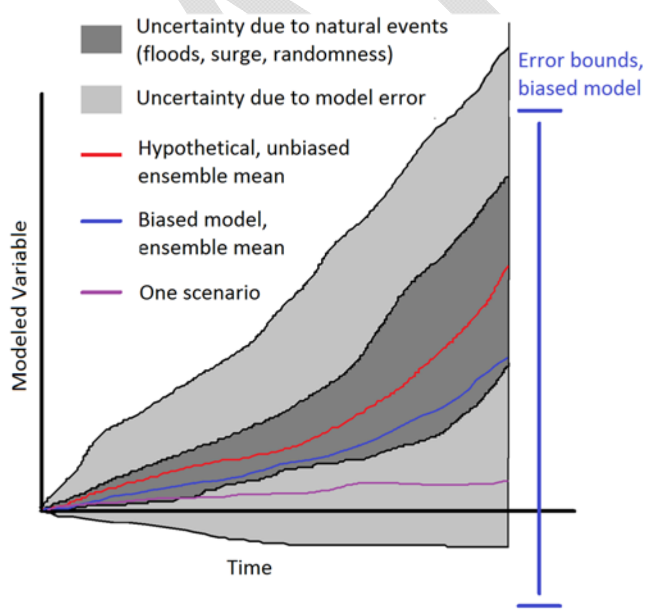

Figure 3: A conceptual argument for ensemble and multi-modeling. Because future environmental conditions are uncertain and models are imperfect, the hypothetical, unbiased ensemble mean (red) is an unreachable goal. It can be approximated by forming an ensemble of outputs that vary model parameters and environmental forcing and produce a biased mean (the blue line). A multi-modeling approach can then be used to reduce bias (not shown). Use of only one simulation with a single model is likely to lead to a prediction (purple) that departs considerably from the ensemble mean due to bias and random errors. (Image courtesy of Dr. Stefan Talke, PSU.)

3. <u>An approach to uncertainty analysis:</u> Summarizing the above, correct evaluation of the impacts of biases and uncertainty propagation is vital to the integrity of the Superfund process. Also, it is best modeling practice to treat propagation of bias/uncertainty through the chained models as an aspect of the model architecture, not as an a posteriori addition [e.g., Reckhow, 1999; Reckhow and Chopra, 1999; Malve et al., 2005; Arhonditsis and Brett; 2004; Borsuk et al., 2004]. This is particularly important in the modeling of ecosystem and biogeochemical processes for which exact equations cannot be formulated, so that any model is, by necessity, statistical or heuristic [Kawamiya, 2002]. Aguilera et al. [2011] recently reviewed the field, described a variety of Bayesian approaches, and provide 76 references to applications in water and water resources. Chen and Pollino [2012] provide additional references and define the current state of best practices.

   To demonstrate what has become "best-practices" over the last 20 years, I quote a meta-review of mechanistic aquatic biogeochemical modeling. With respect to the need for rigorous error analysis of linked models, Arhonditsis et al. [2006] state:

   > *"However, robust modeling tools to address impaired conditions of water bodies are needed now more than ever before; e.g., the costly implementation of total maximum daily loads for pollutants during the next 10-15 years has raised the bar for innovative model developments that can accommodate rigorous error analysis (36). Conceptual weaknesses, methodological omissions, failure to incorporate residual variability, and parameter uncertainty in predictions are more critical when addressing practical management problems (10). In oceanography, the use of models as heuristic tools to elicit conceptual paradigms, to provide semiquantitative (or even qualitative) descriptions and understanding of ecological patterns is still a fundamental objective, while the policy-making process that guides costly management decisions requires predictive tools able to support deterministic statements (and associated errors)."* From Arhonditsis et al. [2006].

   With respect to the importance uncertainty analyses, *Arhonditsis et al.* [2006] state:

   > "*(M)odelers should understand the necessity for explicitly reporting the uncertainty contributed by both model structure and parameters. There is also an urgent research need for novel uncertainty analysis methods that can accommodate complete error analysis and the Bayesian calibration is one of the most promising prospects (38). Bayesian calibration can be used to refine our knowledge of model input parameters, obtain insight into the degree of information the data contain about model inputs (i.e., parameter estimates with measures of uncertainty and correlation among the parameters), and obtain predictions and uncertainty bounds for modeled output variables (44, 45). Technically, this method is a proof*

> *of the concept that there are better ways to parameterize mechanistic models, other than simply tuning (adjusting) model parameters until the modeler obtains a satisfactory fit."* From Arhonditsis et al. [2006].

Regarding use of a Bayesian approach to model calibration, Arhonditsis et al. [2006] continue:

> *"For the purpose of prediction, the Bayesian approach generates a posterior predictive distribution that represents the current estimate of the value of the response variable, taking into account both the uncertainty about the parameters and the uncertainty that remains when the parameters are known (38). Therefore, estimates of the uncertainty of Bayesian model predictions are more realistic (usually larger) than those based on the classical procedures. Predictions are expressed as probability distributions, thereby conveying significantly more information than point estimates in regards to uncertainty (46). The—often deceptive— deterministic statements are avoided and the water quality goals are set by explicitly acknowledging an inevitable risk of non-attainment, the level of which is subject to decisions that reflect different socioeconomic values and environmental concerns."* From Arhonditsis et al. [2006].

The above statements suggest an approach to long-term predictions that would be very useful in the RI/FS process. Rather than providing deterministic predictions from a 45-yr simulation for each remedial alternative, a range of outcomes should be provided that recognize not only accumulated modeling errors and biases, but also uncertainties in future climate and hydrologic conditions.

4. *Recommendations*: Analysis of model errors and biases, and their propagation, should be a fundamental part of the modeling architecture, so that biases and error propagation between models can be evaluated. While a Bayesian approach is not the only possibility, it is presently one of the most widely used in climate modeling. For modeling of long-term outputs Portland Harbor, a multi-modeling hydrodynamic approach could be fairly easily implemented, as the models are not especially computationally intensive, relative, for example, to climate modeling. The 2D model could, for example, be compared to a high resolution 3D model for a limited number of situations, to estimate errors in the 2D model. An ensemble approach to long-term forecasting is also needed, because future environmental conditions are uncertain and influence outcomes. Assumptions regarding both climate change and flood events should be varied. Bayesian techniques should be employed to estimate the overall error/uncertainty in fate and transport predictions due to errors and biases in the hydrodynamic and sediment transport models. If this is not done, then the predictions of the fate and transport models are very difficult to defend.

V. Hydrological Analyses and the 100 year flood

1. Hydrologic analyses and the 100-year flood: As discussed in Appendix I, the 100 year flood in the Willamette is larger than the estimate of the 500-year flood derived in Appendix La of the RI-FS (428,000 cfs). Both the December 1964 and February 1996 floods were measured at about 420,000 cfs (daily average), far larger than the estimated 100-year flood of 360,000 cfs[4]. There have been four floods since 1923 of this magnitude or larger and many more since 1840. The December 1861 flood (the largest since at least 1813) was perhaps 670,000 cfs at Portland, though this value is very poorly constrained.[5] The 1861 flood may represent a flood with a recurrence interval of several hundred years, but fur trader accounts suggest that the flood of 1813 was similar – both flooded Champoeg to about the same level. Clearly, peak flows have been underestimated. The reason why this underestimate occurred is unclear, but use of only about 30+ years of post-1972 data is an obvious factor. On the other hand, WEST Consultants [2004] estimated the 100-year flood at 450,000 cfs, based on the post 1972 data. Thus, it is unclear why a clearly unrealistic estimate of 360,000 cfs was used in the RI/FS.

   Further, the existence of a reservoir system cannot be used as an excuse to ignore historic floods. The Willamette River reservoir system is not especially effective in holding back flood waters, because its capacity is too small, and there are multiple constraints on its management. The threat of a rain-on-snow flood develops quickly, and reservoirs cannot be emptied fast enough to respond. Also, events like the 1861 and 1923 floods emphasize that the reservoir system is not relevant in all floods. Both of these events happened early in the year, before the snow pack was well developed, and both floods were the result of extremely intense and prolonged precipitation. A reservoir system mostly designed to capture Cascade Range snowmelt is largely irrelevant in floods of this sort – most of the precipitation occurs downstream of the dams.

   Finally, the lower Willamette River can also be flooded by backwater effects of the Columbia River. A first pulse of floodwater down the Willamette River might mobilize sediment, which would first be pushed toward the Columbia. It might then be pushed back upriver by backwater flooding from the Columbia. As noted below, there is also no reason to exclude large Columbia River floods from consideration. Despite the reservoir system, an event like 1948 will likely recur, and peak flows in December 1964 were almost as high as those in 1948, though of much shorter duration. The effects of such scenarios on Portland Harbor sediment transport need to be explored.

---

[4] The peak flow measured by USGS in December 1964 was 443,000 cfs.

[5] See http://en.wikipedia.org/wiki/Great_Flood_of_1862 for a summary account of flooding along the West Coast of the US in 1861-1862. See also: http://www.fsl.orst.edu/pnwerc/wrb/Atlas_web_compressed/3.Water_Resources/3e.flood&fema_web.pdf.

2. *Recommendations*: A 100-year flood volume of about 500,000 cfs is realistic and should be adopted. Given climate change and a highly non-stationary Willamette River flow record (see Appendix I), it is difficult to estimate 500-yr flood magnitude. Anything smaller than the 1861 flood is clearly unrealistic. For lack of a better alternative, the 1861 flood could be adopted. Also highly relevant is the issue that the range of past and likely future Willamette River flood styles (and their likely distinct sediment input patterns) should be modeled.

VI. Hydrodynamic Modeling

1. Hydrodynamic Modeling -- Importance: The focus in a Superfund analysis is on contaminants, but contaminant modeling in an estuary or river-estuary context has to be based on hydrodynamic and sediment transport models that have been carefully calibrated and validated. The hydrodynamic model, as the first step in the modeling chain is absolutely critical. If the hydrodynamic model is uncertain or incorrect, all model results based on it are suspect. The Portland Harbor hydrodynamic model does not represent best modeling practices at present, or during the 2004 to 2008 period when the model was implemented. The problem is not the EFDC code, which is widely used. Rather, the problems lie in the manner in which EFDC was implemented. In particular, the grid resolution and area modeled are inadequate to needs of Portland Harbor modeling. Resolution matters because models of sediment and contaminant transport are only as good as the hydrodynamics model that drives then, and sediment/contaminant transport takes place on scales smaller than the scales of variability of tides, water levels and currents. A channel meander of 1m into a contaminated shoal may, for example, release more contamination than decades of slow erosion of broad surfaces. With an underdeveloped grid, there is no way to represent typical fluvial processes that occur on scales of a few meters. The size of the domain modeled matters, because the domain modeled does not have sensible boundary conditions, and the model cannot readily be calibrated and validated in a conventional manner. It should also be mentioned that the evolution of the EFDC hydrodynamic implementation is complex and not well documented. In some cases, it is unclear what aspects of the model discussed in the Phase I and Phase 2 Reports have been retained.

2. Grid resolution and refinement: The model grid cells are typically 25m laterally by 200m alongchannel, with a total of 3355 cells. Multhomah Channel is represented, however, by cells that extend across its width. Baroclinic processes are not represented, and the model is vertically integrated (one layer). This is a very modest grid, even for the 2004 to 2008 period. For example, the baroclinic, unstructured grid model used at Oregon Graduate Institute for nowcast-forecast modeling in the Columbia River estuary and plume in

the early 2000s consisted of 34290 horizontal nodes, 50622 horizontal hybrid elements and 62 vertical levels [Baptista et al, 2005]. While the latter model could fairly be characterized as state-of-the art (and perhaps at a higher level than best-practices for the period), the Portland Harbor grid is far short of best-practices for the period. Among the consequences of inadequate horizontal grid resolution are the following:

a. *Resolution of small-scale features on the bed*: With a very mixed environment in terms of sediment size, complex bed features and erosion pits are likely present. These cannot presently be modeled, and their effects have to be included in the cohesive sediment transport parameterization.

b. *Effects of elongate grid cells*: The aspect ratio of 200m to 25m or 8 is quite large, with associated poor numerical properties. In physical terms, a 200m long grid cell is likely to include quite variable depths and not represent processes well.

c. *Numerical dispersion*: The larger the grid cells, the more numerical dispersion must be included into the model to provide stability. This issue has not been discussed – it should be.

d. *Representation of remedial alternatives*: The present limited grid resolution limits the accuracy of mapping of some remedial alternatives onto the model, decreasing the accuracy of related simulations.

e. *Resolution of structures*: There are numerous bridges across the Lower Willamette that cause locally strong currents and scour, and represent a form of flow resistance. While EFDC does allow their effects to be treated via a bed vegetation algorithm, the limited model resolution means that this cannot be done very accurately.

f. *Analysis of grid resolution*: It is good modeling practice to carry out a grid resolution analysis to determine what grid resolution is required [Roach 1994 and 1997]. Roach [1997] says: "*Systematic grid-convergence studies are the most common, most straightforward and arguably constitute the most reliable technique for the quantification of numerical uncertainty.*" A standard grid resolution analysis would increase grid resolution incrementally until model results converged, and Roach [1997] provides a formalism for determining the convergence rate as the grid is refined.[6] No such tests were carried out as part of Portland Harbor modeling. Also, grid statistics associated with grid aspect ratio, distortion and smoothness should be provided.

[6]While Roach [1994 and 1997] address specifically the sub-category of numerical codes known as "computational fluid dynamics" models, the methodology defined for grid analysis is also valid in the present context.

3. Grid areal coverage: The horizontal extent of the grid in the Columbia River does not allow for imposition of reasonable boundary conditions or testing of the accuracy of water level modeling. For example, backwater effects of the Willamette River on the Columbia extend landward of the model Columbia River grid to Vancouver and beyond, meaning that an elevation boundary cannot readily be applied at the model boundary. There are also not enough tide gauges within the domain for meaningful comparison of observed and predicted water level data. The Columbia River boundaries are also too close to the domain of interest (the Lower Willamette River), possibly introducing errors associated with boundary effects. For purposes of validating model performance and better modeling the complex circulation in Multnomah Channel, the model domain should extend from Willamette Falls to the Columbia River and from Bonneville Dam to at least Longview. This domain encompasses three tide gauges in the Willamette and four in the Columbia. However, two of the four Columbia gauges are at the boundaries (at Bonneville and Longview) and cannot, therefore, be used for calibration. Also, a boundary at Longview is awkward because of the effects of the Cowlitz River. Finally, reproducing the flood of 1996 requires that the domain extend to Beaver, because the Beaver gauge was the only Columbia River gauge that survived during the event.

4. Boundary conditions -- elevation: If elevation boundary conditions are to be imposed, they should be imposed at cross-sections with tide gauges. A larger domain with more gauges in its interior is also needed, as discussed above.

5. Boundary conditions – stream flow: Columbia River flow boundary conditions are poorly described, and their form may have varied over time. However, imposing a flow boundary condition at Vancouver based on flows at The Dalles, as was done at one stage, is unworkable, because the ratio of river flow between The Dalles and Vancouver varies seasonally due to the tributaries between the two locations and because of hydropower operations. Hourly and daily flow values are available for Bonneville Dam since the 1970s (http://www.nwd-wc.usace.army.mil/cgi-bin/dataquery.pl) and should be used as an upstream Columbia River inflow, with the model boundary at Bonneville Dam.

6. Lack of representation of baroclinic processes: Because the hydrodynamic model is 2D rather than 3D, it cannot represent circulation processes related to horizontal and vertical density differences. It is also limited in the realism of its sediment transport calculations in a deep water – often 10-14m in Portland Harbor. Even moderate levels of density stratification strongly affect the vertical structure of velocity and sediment profiles, and horizontal density gradients drive currents. Also, integrated sediment concentrations of ≥2gr/l during the December 1964 flood [Waanenen et al., 1970] are indicative of sediment stratification strong enough to alter vertical mixing, and velocity and sedi-

ment profiles. While the sediment concentrations during this flood are likely a rare circumstance, it is also very important to model such events.

Although 2D models are often employed in rivers, the Lower Willamette is a tidal river, which introduces additional complexity. In particular because of the temperature differences between the Columbia and Willamette and the rapid changes in temperature associated with winter storms, baroclinic effects cannot be *a priori* excluded. WEST Consulants [2004] recognized the need to evaluate whether a 3D model was needed. Phase I modeling [WEST Consulants 2004] says that evaluating this problem is a "primary objective," but then dismisses the issue (in Section 5.1) with a variety of invalid arguments. The conclusions of Section 5.1 also appartently rest on the incorrect idea that the ADCP current profile data in Figures 3.10 and 3.11 of WEST [2004] appear logarithmic. In fact, these profiles represent either defective or inadequately averaged ADCP data from which no firm conclusion can be drawn. One of the profiles Figure 3.10 has (averaging visually) far too much shear above the bed to be representative of a neutrally stratified channel flow. Either the data are defective and should be disregarded, or the conclusion that baroclinic effects are negligible (based on these data) is incorrect.[7]

It is significant that the decision to neglect baroclinic processes is based on discussion with other modelers, but not on any analysis of field data or any comparison whatsoever of 2D and 3D model performance. Combined with the highly superficial model calibration and validation efforts discussed in the next paragraph, the result is an EFDC hydrodynamic model implementation that cannot be relied on.

7. <u>Calibration issues</u>: There has been no systematic comparison of modeled and observed water levels. Comparisons of means is not too the point in a tidal water way. No available report even states what tidal constituents are included in the model. A careful analysis of model's representation of both tides and river stage is needed. In particular, it needs to be demonstrated that tides decrease in the correct manner as flows increase, and that overtides vary in the correct manner with flow. Both tidal constituents and defined water levels (like higher high water HHW, mean water level MWL, and lower low water LLW) should be considered. An example of a simple analysis of HHW that can be used to diagnose processes (when used with data) or model performance (when used with model results) is shown in Figure 4. Methods for analysis of non-stationary river tides and water levels are discussed in Kukulka & Jay [2003a,b], Jay et al. [2011], Matte et al. [2013] and Jay et al. [2014]. These methods should be applied to water level data and model results, to evaluate model performance.

[7] The authors also seem unaware of the fact that a stratified tidal flow has different degrees of density stratification and vertical current variability over the tidal cycle.

No moored ADCP time series data have been used in evaluating model behavior, even though such data have been available since 2003 at the Morrison Street Bridge. The velocity calibration rests on comparisons with lateral profiles on three different days. Again, no systematic comparison of the reproduction of tidal currents and mean flows has been attempted. The same methods described in the previous paragraph for evaluation of water levels can be used to evaluate model reproduction of mean and tidal currents.

8. Validation: Separate calibration and analysis periods are needed to fully validate the EFDC circulation modeling. Each period should be at least a year long and encompass both flood periods and low-flows.

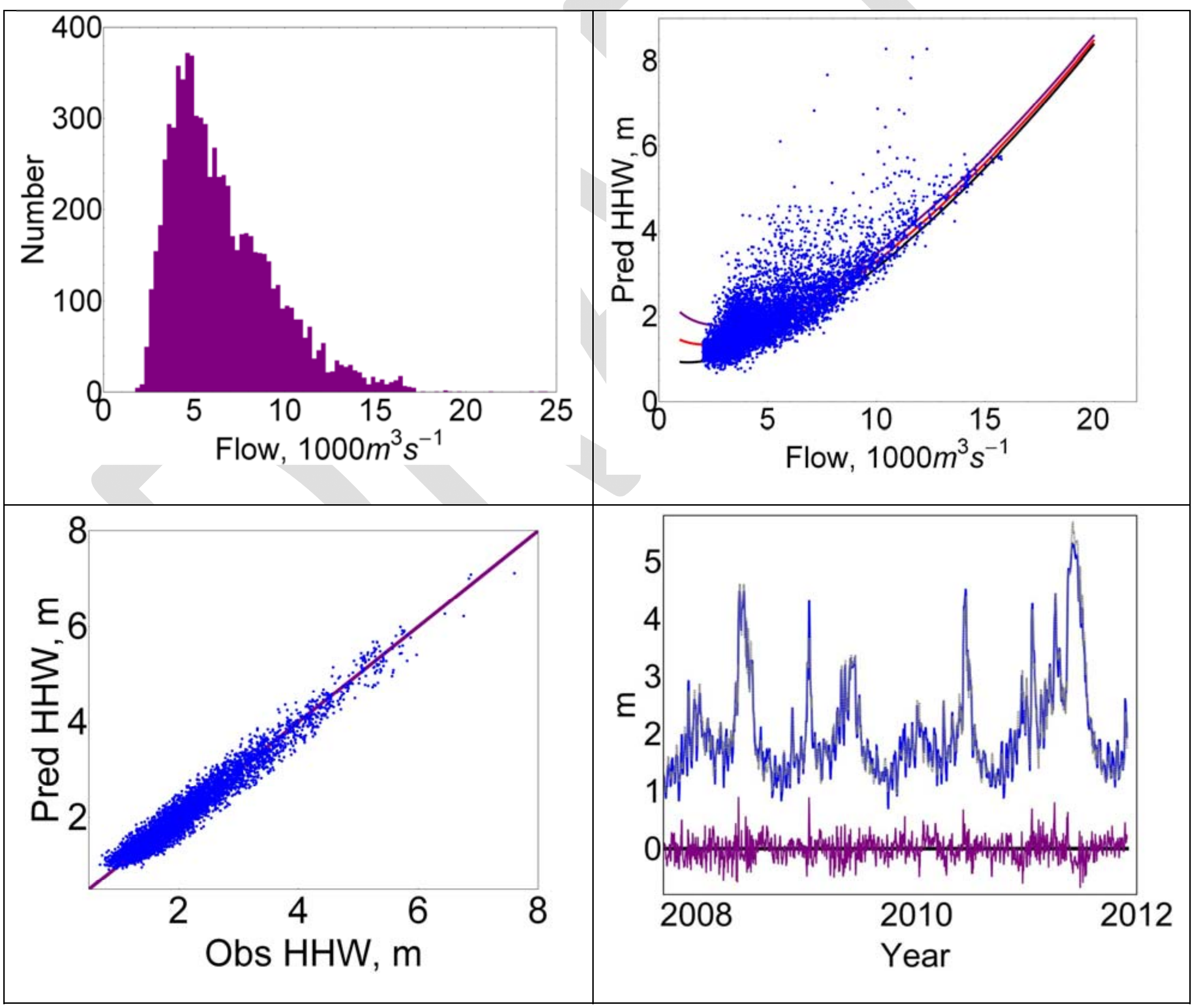

Figure 4: Methods described in Jay et al. [2011] can be used to understand tidal processes and evaluate model performance. Here, observed daily HHW at Vancouver during the 1991-2011 period has been analyzed with a simple regression model that describes the behavior of HHW with 5 parameters: a) 1991-2011 histogram of Bonneville daily flows; b) modeled HHW for three Hammond GDTR values (1.6, 2.6 and 3.6m) and a range of Bonneville flows (Willamette flow =500m^3s^{-1}) on a scatter plot of observed HHW vs. flow; c) hindcast vs. observed HHW, and d) observed HHW (blue), hindcast HHW (grey), and residual HHW (observed - hindcast). The scatter above the model curves in b) reflects the impact of high flows in the Willamette River backwatering the Columbia all the way to Vancouver. As suggested by c), the regression model accounts for >98% of HHW water variance at Vancouver. Analysis of numerical model results in the same manner as the data and comparison of regression model parameters between the tide-gauge data and the numerical model output provide a simple method for analyzing model performance.

9. Sensitivity analysis: The crucial grid-resolution analysis and the absolutely essential determination of whether baroclinic processes need to be included have not been carried out. Once they have been carried out on a revised grid, then a thorough analysis of the effects roughness and turbulence parameters is needed.

10. Bed roughness: Appendix La indicates that bed roughness Z_0 has been set to a constant value, because theoretically based efforts to explain its variability were unsuccessful. While it is better modeling practice to use a constant Z_0 than to use bed roughness as a tuning parameter, the bed in Portland Harbor is quite variable, with areas of fine sediment and sandy sediments with bedforms. Thus, Z_0 should vary. After implementation of a better grid and a careful treatment of other aspects of the hydrodynamic model calibration, this issue should be revisited.

11. Analysis of the 1996 flood: The 1996 flood event was used as a validation exercise. Unfortunately, all Columbia River tide gauges between Longview and Vancouver failed, emphasizing the need to place model boundaries in locations (Beaver and Bonneville Dam) where forcing and validation data are available. Predicted water levels were well above observed water levels, with a discrepancy at the Morrison Street Bridge of up to 0.8m. This indicates unsatisfactory behavior model performance and likely inaccurate predictions of sediment transport. However, few data are available for other parameters aside from water level. The reason for the poor modeling of water levels during the flood is likely related to the grid extent and the boundary conditions used. A better grid will be needed before this issue is re-visited. In some respects, better data are available for the 1964 flood.

12. Recommendations: An elaborate structure of sediment transport and contaminant modeling has been erected on top of an inadequate hydrodynamic model grid and an EFDC implementation that has not been adequately calibrated or validated. Results based on this EFDC implementation cannot be defended. The grid should extend from

Bonneville Dam and Beaver on the Columbia River, and an appropriate grid resolution should be chosen on the basis of model performance and bed properties. This choice of grid will allow appropriate boundary conditions to be imposed. A definitive evaluation of the importance of temperature stratification and horizontal gradients is needed, and this should be based on data that do not now exist; as noted above, these should be collected. A careful calibration and validation of the new grid should be carried out using analytical tools appropriate to the nonstationary nature of the system. The representation of both water levels (tides and stage) and currents (tides and mean flows) should be correct, as both may be important to sediment transport under some conditions. Obtaining a correct representation of these basic processes makes it much more likely that more complex processes like sediment and contaminant transport are being correctly modeled.

VI. Willamette and Columbia River sediment loading

1. Willamette River sediment loading: Sediment supply from the Willamette River is a vital boundary condition for the sediment transport and fate and transport models. Only post-1973 USGS sediment concentration and load data for the Willamette River were used, with observations for days with flows up to about 200,000 dfs. This ignores the larger 1962-1965 daily data set that includes detailed observations for the December 1964 flood, including multiple observations on the days of peak sediment load. The 1964 flood had a peak flow of about 443,000 cfs; it is one of the four largest of the last century. Accordingly, the 1962-1965 data set is an important resource that should have been used. This data set also provides percent sand data, so that the sediment load can be correctly divided into sand and fines transport, and the fines load needs to be divided into silt and clay inputs. Additional problems include:

    a. *Hysteresis effects*: The rating curves derived in Appendix La do not consider sediment load hysteresis, though this is an important factor in the system. Typically, the sediment load is highest on the rising arm of the freshet, and this is an important feature of rain-on-snow floods. Willamette River sediment supply is analyzed in Appendix II.

    b. *Sediment quality*: The RI/FS modeled division of the supply between fines and sand is incorrect for high flows, in part because it does not consider the very large supply of clay material, which is likely most prominent during rain on snow floods.

    c. *Lower Willamette River deposition and erosion*: The sediment load measured at the Morrison Street Bridge does not represent the load to the Lower Willamette

River, because Morrison Street Bridge measurements are affected by deposition and erosion between Oregon City and Portland Harbor. It is likely that the load during low-flow (depositional) periods is underestimated, while the load during high flow periods may be overestimated.

   d. *Use for validation*: The correct use of the Morrison Street Bridge data and rating curve is for validation of the model predictions, not as a boundary condition, because the sampling is within the system, not at the boundary. This problem can only be remedied after collection of an appropriate data set at Oregon City.

2. Columbia River sediment loading: The Columbia River sediment load at Vancouver has been set, based on 1963-1969 data. This is a reasonable first step, but the percent sand has been underestimated. Information in Haushild et al. [1966] should be used to set the percent sand as a function of flow. Also, post 1973 USGS NWIS should be used, as was done for the Morrison Street Bridge.

VII. Sediment transport modeling

Between the hydrodynamic model and contaminant fate and transport modeling there is a sediment transport model (SEDZLJ), which includes modules for water column transport and bed structure. The formulation of the EFDC sediment transport mode appears in most respects to be reasonable. However, the sediment transport modeling is greatly hampered by the use of a 2D horizontal model and the inadequate grid of the hydrodynamic model – the 25 by 200m grid cells are larger than relevant bed features and may encompass areas of variable depth. The treatment of the bed and the active layer between the bed and the flow is also potentially adequate for modeling Portland Harbor. But unless it is actually proven to exist in Portland Harbor and its properties are defined, the active layer should be regarded as a numerically useful abstraction with arbitrary, tunable properties. Moreover, the treatment of the bed and of fine sediment erosion, deposition and settling in the SEDZLJ module of EFDC are, taken together, extremely complex. The number of parameters is sufficient to allow the model to be tuned to correctly represent any particular event, but this does not mean that the model can provide accurate forecasts. Unfortunately, there are few Portland Harbor data available that can be used to objectively set these parameters, and essentially no data for validation of sediment transport predictions. Thus, sediment transport model parameters and results must be regarded as very uncertain, calling into question the fate and transport modeling based on the sediment transport modeling.

There are several specific issues in the sediment transport modeling:

1. Grid resolution: As noted above, the 25 by 200m grid cells are larger than should be used for sediment transport modeling, because they encompass areas of variable depth

and, in some cases, diverse bed properties. Also, the 200:25 aspect ratio may introduce numerical issues in the scalar transport module, in addition to those associated with the performance of the hydrodynamics module. As with the hydrodynamic model, grid resolution test should be run to determine the correct horizontal resolution.

2. Vertically integrated (2D) model formulation: The vertically integrated formulation is even more problematic for sediment transport than it is for hydrodynamics, especially in a deep channel (10-14m in some area) that may also be stratified. It should specifically be demonstrated that a 2D sediment transport model gives the same results as a 3D model, which would certainly be the default choice for the problem at hand.

3. Choice of size classes: Any practical sediment transport model implementation must schematize the continuum of sediment sizes (with its potentially infinite number of sediment classes) into a small number of size or settling velocity classes. Portland Harbor has an extremely diverse range of sediment sizes, from clay to gravel, and floods supply variable amounts of clay, silt and sand. It is also a very short system (<50km) with a limited residence time. During floods, for example, suspended sediment may be advected 10-20km in a day. This makes the distinction between material settling 1m/day and 10m/day critical. In the present implementation of EFDC and SEDZLJ, five size classes have been defined, four for materials that are sands and gravel (fine, medium and coarse sand, and gravel). These four appear appropriate. However, silt and clay (i.e., all materials less than 62.5μ) are treated as a single size class with a settling velocity W_S that varies with concentration and bedstress, but is generally 1-10 m/day. During the December 1964 flood, 33 to 61% percent of four water-column integrated samples taken by USGS during the course of the flood were clay sizes, and 30 to 56% were silt (see discussion in Appendix II). The clay material was mostly washload that did not settle in Portland Habor. But some likely formed flocs and settled, and clay is prominent on the bed in some parts of Portland Harbor. While there are presently insufficient data to make separate rating curves for fines and clay, these two should be distinguished for modeling purposes, because their behavior will be quite different. In summary, a sixth size class is needed, and load and water column size data need to be collected to support this distinction.

4. Settling velocities: The settling velocity formulation for the four size classes of sand and gravel is conventional. For the combined silt and clay size class settling velocity is given by:

$$W_S = 3.3\ (C_1G)^{0.12} \qquad (1)$$

where W_S is in m/day, water column shear stress G is in dyne/cm^2, and C_1 is concentration of size class 1 in mg/l. There are several difficulties with the use of this formulation in the present case:

a) Gradients in shear: Horizontal gradients in shear are high in Portland Harbor, especially during high flow periods. Thus, as a parcel of water moves, the W_S of its load may vary, according to (1). In systems with large spatial scales and slow motions (like most lakes and reservoirs), particles will have time to adjust to their changing environment. This may not be the case in Portland Harbor, and (1) which likely represents equilibrium behavior, may not be appropriate. Unrealistic results may occur both during high-flow periods and in times and places where tidal currents reverse, because shear will change rapidly in both cases.

b) Gradients in concentration: Horizontal gradients in concentration C_1 have not been estimated for Portland Harbor, but the same issue applies to these gradients as to shear gradients. Eq (1) will be unrealistic if the predicted values of W_S change more rapidly (due to advection to a different environment) than the particle field actually responds.

c) Value of G: The shear G is intended to be a water-column value, but bed skin friction shear stress τ_{SF} is used instead, because this is the only value of stress that is available in a 2D horizontal model. If the flow is approximately a channel shear flow, then the shear varies linearly with depth, being maximum at the bed surface and zero at the free surface (unless there is wind). Use of τ_{SF}, which is a component of the bedstress, but not all of it, will mis-estimate the water-column; values may be either too high or too low (if bedforms are present).

d) Problems as slack water: During periods of weak river flow, currents do reverse in Portland Harbor, and slack water is time when sediments typically settle to the bed. The W_S formulation in (1) prevents this from happening by taking W_S to zero as the current slows. This is clearly unrealistic.

5. Absence of modeled bedload transport: As noted above in the CSM discussion, other WLG documents emphasize the important role of bedload transport in bathymetric changes within the Study Area. Thus, it is very surprising to read in Section 2.1 of Appendix La that no bedload transport has been modeled. The reason given is that no formulation of bedload transport over a cohesive bed is available. This is a problem that should be dealt with, given the importance of bedload in the system.

6. Vertically integrated formulation: Use of a vertically integrated formulation inevitably raises difficulties in representation of the four size classes of sand and gravel, whose concentrations are mostly very close to the bed, except during the most extreme flows. Appendix La describes a method to relate the vertically averaged concentration to the near-bed, for the purposes of calculating deposition, which is related to near-bed concentration. However, this approach does not correct the fundamental problem that the transport of a scalar that is strongly concentrated toward the bed is badly overestimates by a vertically integrated model that multiples vertically averaged concentration <C> by vertically averaged velocity <U>. That is, the model estimates Q_S =<U><C> (assuming that the flow is in the x-direction), which is incorrect. That estimate must be corrected by a term that represents the vertical correlations between U and C; thus, Q_S =<U><C>+ <U'C'>, where U'=U- <U> and C'=C-<C>. the variables C' and U' are not calculated by the model. While a theoretical construct (similar to that used to estimate deposition) could be imposed, it does not appear that it is being used. Moreover, a deep river with complex current patterns, factors like lateral shear and density stratification will cause the theoretical construct to fail, at least under some conditions. This is another good argument for use of a vertically integrated model.

7. Coupling of hydrodynamics and sediment transport: EFDC and SEDZLJ are not coupled in the sense that changes in bed elevation (due to deposition and erosion) predicted by SEDZLJ are not coupled back into the EFDC. Under most circumstances, this will not cause major problems in the modeling, and it is a useful simplification for long simulations. However, erosion may reach ~1m during severe flood events. This degree of erosion will change the hydrodynamics. The impacts of this simplification should be judged using fully coupled runs for comparison. Impacts of this simplification also need to be considered in sensitivity analyses.

8. Sediment load time resolution: Given that Willamette River flow can vary by >2000m^3/s, 24-hr period, sediment load can by vary at least by an order of magnitude a day. Thus, sediment load input from upriver should be updated on the same schedule as the river flow – it is unclear whether this is being done at present, and the Phase 2 reports suggest that it is not.

9. Model validation: The validation of the sediment transport model rests entirely on attempts to reproduce observed 2003 to 2009 erosion and deposition patterns, a time period without a major flood. The difficulty with this approach is that it is inherently ambiguous and incomplete. It is impossible to know, even if the bed changes are plausible for the time period, whether the right answer has been reached for the wrong reasons. For example, if a model and data agree that an area shows no net erosion or deposition

over a time period, this does not make the model correct, because erosion and deposition cycles and events that profoundly affect contaminant transport may not have been modeled correctly. Further, the Willamette River sediment load is incorrect (Appendix II) and bedload transport has been neglected. Thus, it is likely that the model's success is based on incorrect parameterizations, calling into question its predictive ability. Given the difficulties documented above in the hydrodynamic and sediment transport models, it is vital that SEDZLJ water column transport predictions be tested against data. While further data collection is needed, there are readily available data sets that have not been used. One is the 2009-2014 USGS time series of turbidity at the Morrison Street Bridge. Acoustic backscatter data or ABS (better for coarser sizes) could also be obtained from USGS for the Morrison Street Bridge side looking ADCP from 2003-2014. Both time series should be calibrated, considering variations in both particle size and concentration.

10. Data collection needs: Additional data collection should include:

    a. Concentration data from turbidity: Time series of sediment concentration data (from turbidity) are needed from multiple levels at at least two locations, e.g., the Morrison Street Bridge and St Johns Bridge, for at least one year, preferably more. Turbidity should be calibrated to concentration using the methods described in USGS [2009].

    b. Concentration data from ABS: As discussed previously, moored ADCP data are needed to calibrate the EFDC hydrodynamics module on an improved grid. The ADCP deployments will yield ABS data that can be used to determine sediment concentrations.

    c. Moored LISST data: Time series of size data should be collected so that the sizes represented by the turbidity data can be determined under a wide variety of conditions.

    d. Water samples: laboratory analyses of disaggregated size and organic content are also need, to calibrate the turbidity and ABS data from existing and future time series.

    e. Sediment load at Oregon City: The sediment load at the upstream boundaries of system at Oregon City and from the Clackamas river need to be determined, including both quantity and quality. This will require several years of measurements using the methods described above.

11. The bed model in SEDZLJ: As noted, the bed model employed is complex. In the end, horizontally uniform erosion properties were specified for cohesive sediments, because the 15 cores available for Sedflume analyses were not sufficient to define spatially variable properties. This is not *per se*, a bad idea. However, it may interact badly with the inadequate grid resolution and vertically integrated model formulation. The sensitivity analyses need to be carried out in a 3D model with better grid resolution to determine how to best represent bed erosion properties.

12. Sensitivity analyses and error propagation: The efforts made to date to test the sensitivity of the SEDZLJ module are commendable but incomplete, in that a) error propagation from the hydrodynamics model and its boundary conditions is not recognized; b) errors associated with the vertically integrated formulation are not considered; and c) the number of parameters tested is quite limited. Moreover, the decision to discard the highest and lowest results for the 26 model runs in examining errors in the fate and transport modeling is probably not justified. A Bayesian framework would consider the entire spectrum of outputs in a probabilistic manner, and provide a more realistic result.

13. Modeling of the 1996 flood: Attempts to verify the sediment transport modeling using the 1996 flood are incomplete, because there two data points that can be used to verify water column sediment concentrations and no data to test the predicted patterns of deposition and erosion. It is possible that historic bathymetric data before and after the event could be recovered from the Port of Portland or the US Army Corps of Engineers. Also, as discussed in Appendix II, the sediment loading function used for the Willamette River is incorrect, at least for high flow conditions. Finally, Appendix La suggests that the 1996 flood approaches the 500-yr level. This is incorrect – it is only the third or fourth largest flood in the last 90 years.

14. Recommendations: The validity of the 2D horizontal sediment transport formulation needs to be demonstrated on a more detailed and extended grid than used at present, and numerous details of the model formulation need further consideration. Bedload transport should be included, clay and silt need to be represented as separate size classes, and sediment input from the Willamette River should be modeled correctly, requiring data collection at Oregon City. The model needs to be calibrated with and validated against more extensive field data (to be collected) that allow the predicted transports to be tested. Further sensitivity analyses error estimates are needed that include error propagation from the EFDC hydrodynamic module.

VIII. Fate and transport modeling

The general approach to scenario analysis used in the fate and transport modeling is appropriate, and QEAFATE is a useful tool, if properly verified. But fate and transport modeling can only be as accurate as the hydrodynamic and sediment transport modeling that supports it, because errors and uncertainty propagate and grow as modeled are chained together in succession. As such, the fate and transport results in Appendix Ha of the RI/FS should be regarded as highly preliminary. Moreover, there are also issues specific to implementation of the fate and transport module QEAFATE, as described in Appendix Ha:

1. <u>Boundary loading of contaminants</u>: Even though most contaminants modeled are strongly associated with particulates (and to some extent, dissolved organic matter), contaminant loading is specified as a function of flow, rather than input suspended sediment concentration. The latter would be more logical. In any event, contaminant loading is likely also susceptible to hysteresis effects that should be defined.
2. <u>Mapping of scenarios onto the grid</u>: Section 5.3.1.1 of Appendix Ha indicates that the grid is too coarse to accurately map the remedial alternatives onto the bed. While this may not be the most serious problem associated with limited grid resolution, it is one of the issues.
3. <u>Calibration period</u>: The post 2003 calibration period does not have a really major flow event. While available data may require this period to be used, this still represents a limitation on the model capabilities.
4. <u>Sensitivity analyses</u>: The uncertainty analysis does recognize the importance of sediment loading, but no other sources of uncertainty and bias associated with the hydrodynamic and sediment transport modeling are recognized. The result is that uncertainties are far higher than reported. A bayesian approach to uncertainty is likely the best way to approach the uncertainty analysis.
5. <u>100-year flood event</u>: The hydrologic forcing for the 45-year model runs includes the 1996 flood, which is appropriate. However, the 1996 flood (maximum daily flow of about 420,000 cfs) was 16% short of a 100-year flood event (about 500,000 cfs; Appendix I). If impacts of a 100-year flood are to be analyzed, additional scenarios will need to be analyzed.
6. <u>Variability of natural conditions</u>: For each remedial alternative, a spectrum of forcing scenarios should be considered, to determine how variable future outcomes may be, including effects of long-term changes in hydrology and climate.
7. <u>Bed layering</u>: Bed layering is different in QEAFATE and SEDZLJ. This raises issues of model consistency that should be explained.

8. Recommendations: Fate and transport modeling needs to be based on improved hydrodynamic and sediment transport modeling, multiple scenarios should be recognized for each remedial alternative, and a proper framework for analyses of error and uncertainty is needed.

IX. Recommended future analyses

There are a number of steps that EPA could take with existing data and models that would improve the understanding of the Portland Harbor Superfund site, its history, and relevant processes, and provide a better ability to evaluate model outcomes. These include:

1. Historical data: Several tide gauges were deployed in the Columbia River in the 1960s [Yeh et al., 2012], and gauges (other than the Morrison Street Bridge gauge) may have been deployed in Portland Harbor. This suggests that that the 1964 flood, with its valuable data set may be susceptible to modeling. However, the "data archaeology" first needs to be done to rediscover and recover the historical data [cf. Talke & Jay, 2013]. It is also possible that historic bathymetric data for 1996 could be recovered from the Corps of Engineers and/or the Port of Portland.
2. Water level data analyses: Water level data are the simplest physical data to work with and reveal a great deal about a system, but their analysis has been completely ignored in Portland Harbor Superfund project. Water level data exist for Portland Harbor and Vancouver back to the 1870s, with several years of tide data from two stations in Portland Harbor between 1900 and 1915. There are extensive hourly data (three stations in Portland Harbor plus Vancouver and St Helens) for 1941 to 1943, and three station with hourly data almost continuously since the late 1980s or early 1990s, along with water level data from below Bonneville, Vancouver and St Helens. Thus, the data exist to provide very precise information (via non-stationary water level analysis methods; Jay and Flinchem, 1997; Kukulka nd Jay 2003a,b; Jay et al., 2011; Jay et al., 2014) about how water levels in the harbor respond to river flow, and how tidal range varies with flow and the neap-spring cycle. Long-term trends in water level properties can also be determined, which is important for projecting future conditions. As noted above, the analyses we apply to data can also be applied to numerical model results, to see if the model responds correctly. Thus, non-stationary water level analysis methods are a very important tool for numerical model calibration, but have not been used here.
3. Analysis of USGS data sets related to turbidity: USGS has collected time series of turbidity (2009-2014) and ADCP backscatter (ABS, 2003-2014). These could be analyzed to better understand suspended sediment parameters. NWIS samples (the same data set used to establish sediment load rating curves) could perhaps be used for calibration. It is

likely, however, that this sampling is not adequate for this purpose, and further data collection is needed. A preliminary analysis is, nonetheless, possible.

4. Bed sediment and sedimentology: The GeoSea [2001] data set and analyses, and subsequent bed data sets should be examined carefully, to provide a better qualitative understanding of Portland Harbor sedimentation and sediment transport processes. It is important to determine how coarse material present on the bed was emplaced, whether it is presently mobile, and whether the sediment transport model is working correctly. Sedimentology can also give important insights into past and likely future extreme events.

5. Hydrologic analyses: The “design flood” for the Willamette River almost inevitably involves interactions between Columbia and Willamette River flows. The greatest bed-stresses are likely to be produced, not by the most extreme high waters or high flows, but during periods of rapid increases in flow from a relatively low water level. Also, it is not reasonable to project 100 and 500 year floods on the basis of 30-40 year flow records, when daily flow measurements date back to 1878 in both the Willamette and Columbia. Determining realistic extreme events and trends therein requires use of these records and consideration of long-term trends in flow properties of both systems – flood risk is not stationary; it is evolving due to climate change and human management. There are two likely approaches:
    a. *Analyses of existing data*: Although we have produced routed flows to Portland Harbor for 1878 to date (Jay and Naik, 2011), these estimates could be improved in a number of ways, to improve the estimates of peak flows.
    b. *Climate and hydrologic modeling*: Recent advances in downscaling of global climate model results to provide regional projections could also be used to advantage. Groups at Portland State University and Oregon State University are conducting extensive analyses of possible future flood and drought conditions in the Willamette Valley, and the results of these should be used in the superfund process.

6. Sediment load analyses: The results of Appendix II form a good “first look” at the Willamette River sediment load problem, but much more could be done. I did not analyze the City of Portland data set, because these data were not available to me. There are also suspended sediment data collected by the US Army Corps of Engineers for several years around 1950. There are also likely data from the early 1920s, though the only data set I have seen pertains to the Columbia River. Even if there are no data, it would be very useful to analyze the 1923 flood from contemporary accounts. Also, Corps of Engineers dredging records should be examined to determine historic sediment accu-

mulation. This sort of historical analysis would also be quite useful in understanding what future floods in the system will be like, in the 2040-2060 period, at the end of the monitoring period after remediation.

7. Analyses of water levels and bedstresses: The maximum potential for erosion of contaminants during a flow event occurs when bedstresses are high, which does not necessarily correspond to the peak of the flood. Existing water level time series and numerical models can be used to determine the conditions leading to maximum bedstress in Portland Harbor, though moored ADCP would be very helpful in such a study.

8. Remote sensing of turbidity: Remote sensing (ocean color products like MODIS) has long been used to analyze surface turbidity patterns in open coastal waters. These methods have recently been demonstrated to work well in the Columbia River estuary [Hudson, 2014]. The coarse spatial resolution of MODIS (250m) precludes its use in Portland Harbor. Free Landsat imagery (30-60m resolution, 16day return period) could, however, be used, in combination with the USGS Morrison Street Bridge turbidity and ABS time series, to analyze spatial patterns of turbidity in Portland Harbor and the Lower Willamette more generally, including at Oregon City. Because fine sediments are nearly uniform in the vertical, such an analysis would be very helpful in understanding turbidity patterns in Portland Harbor, and could also be used for model verification, as Landsat data have been collected since the 1970s. Landsat 5 (1984-2013), Landsat 7 (beginning in 1999), and Landsat 8 (since 2013) data would be the most relevant Landsat missions. Also, high-resolution images from private multi-spectral satellites were de-classified in June 2014. These images have a resolution of 1-5m, ideal for fluvial studies. These satellites missions began between 1999 and 2009, depending on the satellite. While these images remain expensive for universities and private contractors, government agencies may be able to obtain them at cheaper rates. PSU has considerable experience in analysis of turbidity maxima processes and remote sensing of turbidity.

9. Multi-modeling: It is important to use existing models to evaluate Superfund modeling. For example, Portland State University has a more detailed (in terms of grid resolution) numerical model of Portland Harbor and the lower Columbia River, implemented in the Delft3D system. USGS (Portland District) also used uses Delft3D for its sediment transport modeling. Our grid covers the entire system (including floodplain) from Bonneville Dam and Oregon City to the ocean. It is 3D below Beaver, and could easily be made 3D in Portland Harbor. Like EFDC, Delft3D has sediment and contaminant transport modules that build on a hydrodynamic module, but we have run only the circulation module so far. Use of this more detailed model would allow a more careful

analysis of water level, current and bedstress patterns than is possible with the implementation of EFDC used in the RI/FS.

References

Aguilera, P.A., Fernández, R., Fernández, A., Rumi, R., Salmerón, A., 2011. Bayesian networks in environmental modelling. *Environ. Model. Softw*. 26., 1376e1388.

Arhonditsis, G.B., Adams-VanHarn, B.A., Nielsen, L., Stow, C.A., Reckhow, K.H., 2006. Evaluation of the current state of mechanistic aquatic biogeochemical modeling: citation analysis and future perspectives. *Environ. Sci. Technol.* 40, 6547–6554.

Arhonditsis. G.B.. Brett, M.T.. 2004. Evaluation of the current state of mechanistic aquatic biogcochemical modeling. Marine Ecology Progress Series 271. 13-26.

Best, J., 2005, The fluid dynamics of river dunes: A review and some future research directions, *J. Geophys. Res*., 110, F04S02, doi:10.1029/2004JF000218.

Borsuk, M. A., C. A. Stow, and K. H. Reckhow. 2004. A Bayesian network of eutrophication models for synthesis, prediction, and uncertainty analysis. *Ecolog. Model.* 173:219–239.

Brands, M. D. (1947) Flood runoff in the Willamette Valley, Oregon. *USGS Water Supply Paper*, 968A, 59pp.

Chen, C.F., Ma, H.W., Reckhow, K.H., 2007. Assessment of water quality management with a systematic qualitative uncertainty analysis. *Sci. Total Environ*. 374, 13–25.

Chen, S. H. C. A. Pollino, 2012. Good practice in Bayesian network modelling, *Environ. Model. Softw*., 37, 134-145.

Deems, J., Hamlet, A. F. (2010). Historical meteorological driving data set. *Final Report for the Columbia Basin Climate Change Scenarios Project*, Climate Impacts Group, UW, Seattle.

Dettinger, M. D, (2011). Climate change, atmospheric rivers, and floods in California – A multi-model analysis of storm frequency and magnitude changes. *Journal of the American Water Resources Association 47*(3), 514-523. DOI: 10.1111/j.1752-1688.2011.00546.x.

Haushild, W. L., R. W. Perkins, H. H. Stevens, G. R. Dempster, and J. L. Glenn, 1966. *Progress report: radionuclide transport in the Pasco to Vancouver, Washington reach of the Columbia River July 1962 to September 1963*. Portland, Oregon, U. S. Geological Survey. 188 p. (Prepared in co-operation with the U. S. Atomic Energy Commission.)

Jay, D. A. and Flinchem, E. P., 1997, Interaction of fluctuating river flow with a barotropic tide: A test of wavelet tidal analysis methods, *J. Geophys. Res.* **102**: 5705-5720.

Jay, D. A., B. S. Giese, and C. R. Sherwood, 1990, Energetics and sedimentary processes in the Columbia River Estuary. *Prog. Oceanogr.*, **25**: 157-174.

Jay, D. A., K. Leffler and S. Degens, 2011, Long-term evolution of Columbia River tides, *ASCE Journal of Waterway, Port, Coastal, and Ocean Engineering*, 137: 182-191; DOI: 10.1061/(ASCE)WW.1943- 5460.0000082.

Jay, D. A., K. Leffler, H. Diefenderfer, and A. Borde, 2014, Tidal-fluvial and estuarine processes in the Lower Columbia River: I: along-channel water level variations, Pacific Ocean to Bonneville Dam, *Estuaries and Coasts*, DOI 10.1007/s12237-014-9819-0.

Jay, D. A., and P. Naik, 2011, Distinguishing human and climate influences on hydrological disturbance processes in the Columbia River, USA, *Hydrological Sciences Journal,* 56, 1186-1209.

Kawamiya, M., 2002, Numerical model approaches to address recent problems on pelagic ecosystems. *J. Oceanogr.* 58, 365- 378.

Kukulka, T., and D. A. Jay, 2003a, Impacts of Columbia River discharge on salmonid habitat I. a non-stationary fluvial tide model, *J. Geophys. Res*. **108**, 3293 doi 10.1029/2002JC001382.

Kukulka, T., and D. A. Jay, 2003b, Impacts of Columbia River discharge on salmonid habitat II. Changes in shallow-water habitat, *J. Geophys. Res*. **108**, 3294 doi 10.1029/2003JC001829.

Lower Willamette Group (LWG), 2006. Portland Harbor RI/FS Phase 2 Recalibration Results: Hydrodynamic Sedimentation Modeling for the Lower Willamette River Draft, prepared by WEST Consultants and Integral Consulting, Inc.

Malve, O., M. Laine, H. Haario, 2005. Estimation of winter respiration rates and prediction of oxygen regime in a lake using Bayesian inference, *Ecolog. Model*. 182, 183–197.

Matte, P., D. A. Jay and E. Zaron, 2013, Adaptation of Classical Tidal Harmonic Analysis to Non-Stationary Tides, with Application to river tides, *J. Ocean. Atmos. Technol*. 30: 569-589, doi 10.1175/ JTECH-D-12-00016.1.

Moftakhari, H. R., D. A. Jay, S. A. Talke, T. Kukulka, and P. D. Bromirski, 2013, A Novel Approach to Flow Estimation in Tidal Rivers, *Water Resources Research*, 49, doi:10.1002/wrcr.20363.

Naik, P.K., and D.A. Jay, 2005, Virgin flow estimation of the Columbia River (1879-1928), *Hydrologic Processes*, 10.1002/hyp.5636.

Naik, P. and D.A. Jay. 2011. Human and climate impacts on Columbia River hydrology and salmonids. *River Research and Applications* 27: 1270-1276 DOI: 10.1002/rra.1422.

Nelson, E. (1949). Columbia River Basin flood May-June 1948. *Monthly Weather Review, 77*, 1-10.

Paulsen, C. G. (1949). Floods of May-June 1948 in Columbia River Basin, *USGS Water Supply Paper 1080*, Washington.

Reichert, P., Omlin, M., 1997. On the usefulness of overparameterized ecological models. *Ecol. Model.* 95, 289–299.

Reckhow K. H., 1999. Water quality prediction and probability network models. *Can. J. Fish. Aquat. Sci.* 56, 1150–1158.

Reckhow K. H., and Chapra S. C., 1999. Modeling excessive nutrient loading in the environment, *Environ. Pollut.* 100 197-207.

Speers, D. D. D. J. Barcellos and R. T. Wortman. 1990. The 1948 flood on the Columbia River an analysis of its cause and its importance in operational planning, Western Snow Conference, 1990, Sacramento.

Roach, P. J., 1994. Perspective: a method for uniform reporting of grid refinement studies, *J. Fluids Engin.* 1994: 405-413.

Roach, P. J., 1997. Quantification of uncertainty in computational fluid dynamics, *Annu. Rev. Fluid. Mech.* 29, 123–60.

Talke, S. A. and D. A. Jay, 2013, Nineteenth Century North American and Pacific Tidal Data: Lost or Just Forgotten? *J. Coastal Res*., doi: http://dx.doi.org/10.2112/ JCOASTRES-D-12-00181.1.

Templeton, W. and D. A. Jay, 2013, Lower Columbia River sand supply and removal: estimates of two sand budget components, *ASCE Journal of Waterway, Port, Coastal, and Ocean Engineering*, doi:10.1061/(ASCE)WW.1943-5460.0000188.

Sherwood, C. R., D. A. Jay, R. B. Harvey, P. Hamilton and C. A. Simenstad, 1990. Historical changes in the Columbia River estuary. *Progr. Oceanogr*. **25**: 271-297.

USGS, 2009. Guidelines and Procedures for Computing Time-Series Suspended-Sediment Concentrations and Loads from In-Stream Turbidity-Sensor and Streamflow Data, Chapter 4 of Book 3, Applications of Hydraulics Section C, Sediment and Erosion Techniques. Techniques and Methods 3–C4, US Department of Interior, USG Geological Survey, Reston, VA.

Yeh, H., E. Tolkova, D. A. Jay, S. A. Talke, H. Fritz, 2012, Tsunami Hydrodynamics in the Columbia River, *Journal of Disaster Research* 7: 604-608.

Zhang, W., G.B. Arhonditsis, 2009. A Bayesian hierarchical framework for calibrating aquatic biogeochemical models *Ecological Modelling* 220, 2142–2161.